EVALUATION OF GREEN STORMWATER INFRASTRUCTURE MONITORING PROTOCOLS

Lauren M. Cetin

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Civil Engineering

Randel L. Dymond, Chair
Clayton C. Hodges,
Kevin D. Young

May 8, 2018
Blacksburg, VA

Keywords: stormwater, monitoring, green stormwater infrastructure, best management practice, monitoring protocol
Evaluation of Green Stormwater Infrastructure Monitoring Protocols

Lauren M. Cetin

ACADEMIC ABSTRACT

To address and counteract impairment of waterways in the United States, best management practices (BMPs) or Green Stormwater Infrastructure (GSI) have been extensively implemented without any long-term monitoring studies completed; instead relying on lab-tested or assumed pollutant removal efficiencies that often do not translate into field implementation. Monitoring studies have often been applied with variable standards, which lead to inconsistent results and inconclusive data. This study aims to synthesize essential components of a GSI monitoring program based on a review of existing national (Technology Assessment Protocol – Ecology [TAPE], Technology Assessment Reciprocity Partnership [TARP], etc.) and local monitoring standards. Data from past protocols was used in tandem with historic precipitation data to develop a methodology for creating a local or small region-specific protocol. This methodology was applied to the case study area of Fairfax, Virginia. Results from the study indicate that historic precipitation data and past protocol recommendations can be effectively applied in a local setting to create a more suitable protocol adapted for GSI monitoring.
Evaluation of Green Stormwater Infrastructure Monitoring Protocols

Lauren M. Cetin

GENERAL AUDIENCE ABSTRACT

Due to development of once natural landscapes, also referred to as urbanization, stormwater management has evolved in an effort to address and counteract impairment of waterways in the United States by extensively implementing best management practices (BMPs) or Green Stormwater Infrastructure (GSI). Facilities are installed without any requirement of long-term monitoring; instead relying on lab-tested or assumed pollutant removal efficiencies that often do not translate into field implementation and do not perform as intended and required by regulatory agencies. Monitoring studies have often been applied with variable standards, which lead to inconsistent results and inconclusive data. This study aims to synthesize essential components of a GSI monitoring program based on a review of existing programs (Technology Assessment Protocol – Ecology [TAPE], Technology Assessment Reciprocity Partnership [TARP], etc.). Data from past protocols was used in tandem with historic precipitation data to develop a methodology for creating a local or small region-specific protocol. This methodology was applied to the case study area of Fairfax, Virginia. Results from the study indicate that historic precipitation data and past protocol recommendations can be effectively applied in a local setting to create a more suitable protocol adapted for GSI monitoring in order to confirm designed efficiency.
ACKNOWLEDGEMENTS

The completion of this thesis would not have been possible without the support of many individuals. First and foremost, I would like to thank my committee chair and advisor, Dr. Randel Dymond, for providing guidance, support, and constantly pushing me to achieve more. I would also like to extend my sincere appreciation and gratitude to Dr. Clayton Hodges for his time, wisdom, and abnormally efficient work-ethic. Clay has been a tremendous help throughout this entire process, constantly assuring me that “it is going to be okay” and always being available to talk through a question day or night. I would also like to thank Kevin Young for being a member of my committee and being the one to tempt me with the idea of pursuing a Master’s Degree. Kevin has guided me in becoming a better teacher, professional, and person, and for that I am extremely thankful. My committee was an invaluable part of my experience and I can honestly say I would not have made it to where I am today without them.

I would also like to thank my fellow research group members, Amanda, Meghan, Melissa, and Zack, along with the newest additions of Conrad, David, Dylan, Paul and Zoe. This research group was like a family to me, providing hugs, laughs, and advice whenever needed.

Lastly I would like to thank my friends and family for their continued support and understanding throughout these extremely stressful times, as well as an apology to my cat, Keiko, for all the long days and nights I spent in the office and not at home.
# Table of Contents

Academic Abstract ............................................................................................................. ii
General Audience Abstract ............................................................................................ iii
Acknowledgements ........................................................................................................... iv
Table of Contents ............................................................................................................. v
List of Figures .................................................................................................................... vii
List of Tables ..................................................................................................................... viii

1. Introduction .................................................................................................................. 1
   1.1 Background ........................................................................................................... 1
   1.2 Problem Statement ............................................................................................... 2
   1.3 Purpose and objectives ......................................................................................... 2

2. Literature Review ........................................................................................................ 4
   2.1 Impacts of urbanization ......................................................................................... 4
   2.2 Evolution of Stormwater Management Regulations ............................................... 4
   2.3 Evolution of Current Stormwater Management ....................................................... 5
   2.4 The importance of Monitoring ............................................................................ 6
   2.5 Existing Monitoring Protocols ............................................................................. 6
       2.5.1 Environmental Technology Verification Program (ETV) ................................ 6
       2.5.2 Technology Acceptance Reciprocity Partnership (TARP) .............................. 7
       2.5.3 Technology Assessment Protocol – Ecology (TAPE) ..................................... 8
       2.5.4 Virginia Technology Assessment Protocol (VTAP) and Chesapeake Bay...
           technology Assessment Protocol (CBTAP) .......................................................... 9
   2.6 Notable Monitoring parameters .......................................................................... 9
       2.6.1 Minimum Number of Storms Sampled ........................................................... 9
       2.6.2 Inter-event Time ............................................................................................ 10
   2.7 Conclusion ............................................................................................................. 11

   3.1 Introduction .......................................................................................................... 12
       3.1.1 Background .................................................................................................. 12
       3.1.2 Existing Protocols ....................................................................................... 14
       3.1.3 Purpose ..................................................................................................... 20
   3.2 Methods ............................................................................................................. 21
       3.2.1 Analysis of Existing Protocol Components ................................................... 21
       3.2.2 Historic Precipitation Analysis ................................................................ 24
3.3  Case Study: Development of Monitoring Protocol for Fairfax County, Virginia .......25
3.4  Required Monitoring Components .................................................................25
3.5  Storm Event Components ..............................................................................26
3.6  Results and Discussion ..................................................................................31
   3.6.1 Assignment of Protocol Components by Weighted Dichotomous Data Scoring ....31
   3.6.2 Application of Automated Local Precipitation Analysis ..............................31
3.7  Conclusion ........................................................................................................32
4.   Conclusion .........................................................................................................34
   4.1  Implications ...................................................................................................34
   4.2  Future Work ...................................................................................................34
5.   References .........................................................................................................36
LIST OF FIGURES

Figure 3.1. Geographic location of all monitoring programs collected for this study, including place of origin as well as recognizing states or cities ................................................................. 16

Figure 3.2. Flowchart depicting typical items associated with creating a monitoring program .... 20

Figure 3.3. Location Map of Reviewed Precipitation Gauges with Gauge ID noted ............ 27

Figure 3.4. Average number of qualifying storm events for consolidated data from all gaging stations in Fairfax County, VA based on a minimum 0.25 cm storm depth and 6 hour inter-event time .................................................................................................................. 28

Figure 3.5. Average number of qualifying storm events for consolidated data from all gaging stations in Fairfax County, VA based on a minimum 1.27 cm storm depth and 6 hour inter-event time .................................................................................................................. 30
LIST OF TABLES

Table 3.1. Summary Table of Reviewed Monitoring Protocols .......................................................... 15

Table 3.2. Quantitative Comparison of Existing Protocols. NCPEP and MASTEP were not included in this table due to their lack of protocol documentation available. ....................... 22

Table 3.3. Qualitative Comparison Summary Table of Existing Protocols: Type of facility – “N”: non-proprietary, “P”: proprietary, “B”: both. An entry of 1 and 0 correspond to a yes or no, respectively. NCPEP and MASTEP were not included in this table due to their lack of protocol documentation available. .......................................................................................... 23

Table 3.4. Summary Table of Recommended Storm Event Parameters for Fairfax County, VA. 30
1. INTRODUCTION

1.1 BACKGROUND

Current structural stormwater best management practices (BMPs) are implemented to mitigate quantity and quality of urban runoff on a spatially distributed, site scale. These BMPs are constructed to capture, treat and slowly release or infiltrate on-site flow, in the attempt to mimic pre-development conditions and thereby mitigate downstream problems such as sediment and nutrient loading in order to meet region specific total daily maximum loads (TMDLs) (Hopkins et al. 2017; Ice 2004; Livingston at al. 1997; Sample et al. 2012). These facilities are also referred to as green stormwater infrastructure (GSI), and defined specifically as quality-driven stormwater control measures (SCMs). The increased reliance on these facilities to comply with regional and local stormwater regulations begs the need for confirmation of treatment efficiency through consistent long-term monitoring programs. Monitoring programs have been implemented at the national-, regional-, state-, and city level with varying level of detail, and varying levels of success due to either funding, lack of regulatory backing, or lack of interest (Adair et al. 2014).

Although the specifics of each monitoring program vary, they all are comparable in terms of structure of design and implementation. The core steps of any monitoring plan are defining objectives, establishing parameters of interest, planning to collect data using specific sampling frequencies, developing an analytical approach of data, and redefining a plan as necessary (Geosyntec Consultants 2009). Parameters of interest vary based on region and watershed characteristics, however, a common issue in waterways is an abundance of sediment. Total Suspended Sediment (TSS) is not just an indicator of upstream erosion and surface runoff but also can act as a proxy for estimation of other pollutants that bond easily with sediment such as metals (DDOE 2014). Despite varying needs and problems within individual municipal separate storm sewer systems (MS4s), a standard set of methods for GSI monitoring that can be easily adapted would be beneficial for the long-term assessment of facility efficiency.
1.2 PROBLEM STATEMENT

The process of creating a monitoring plan as well as obtaining and analyzing meaningful data is not well documented in literature or practice. There are various national-, state-, and city-scale protocols developed for providing guidance on monitoring stormwater treatment facilities. Methods for selecting monitoring parameters such as storm event sample size, or minimum amount of rainfall vary between protocols with no explanation or defined reasoning for selection. Due to these gaps, a uniform document would benefit the widespread implementation and confidence in GSI.

1.3 PURPOSE AND OBJECTIVES

While the typical components necessary for implementation of monitoring plans is consistent across protocols, there has been little guidance regarding adaptation of historic protocols for specific regional application. Many of the components, such as equipment, lab analysis procedures, and data analysis are standard and are directly translatable across large regions. However, development of storm event parameters should include a detailed analysis of local data to design the most comprehensive monitoring plan. A methodology to identify essential components of a monitoring protocols and determine monitoring and sampling frequencies is necessary to design consistent translatable monitoring efforts that will aid the scientific community in more fully understanding the functionality of GSI. The purpose of this study is to identify essential components of a monitoring protocol for non-proprietary stormwater facilities and provide a methodology for analysis of precipitation data for identification of many protocol parameters, which was implemented on a case study in Fairfax County, Virginia. The objectives of this study are as follows:

1) Review existing literature on monitoring GSI facilities, protocols and analyzing data
2) Analyze and compare existing protocols of monitoring programs such as TAPE, TARP and ETV to better understand decision making process of measured quality parameters
3) Provide guidance on effective sampling and monitoring frequencies of conducting monitoring studies based on geographic region
Although this study focuses on non-proprietary GSI facilities, a majority of past protocols have been designed specifically for proprietary manufactured treatment devices (MTDs), which are often low-volume, low-flow stormwater quality facilities with smaller contributing drainage areas (CDAs). Despite this difference, the overall monitoring procedures are similar with special consideration taken into qualifying storm details such as minimum storm depth, which would vary for proprietary vs. non-proprietary facility type.
2. **Literature Review**

2.1 **Impacts of Urbanization**

The most enveloping and rapid form of land use conversion in the United States is the transition from agricultural or forested land to urban land, commonly referred to as “urbanization” (Schoonover and Lockaby 2006). Urbanization is characterized by increased impervious surfaces, and replacement of natural drainage systems with networks of storm pipes and channels (Nelson et al. 2006). Urbanization leads to impacts on watershed hydrology such as increased peak flows, increased runoff volumes, increased sediment loading and overall degradation of water quality (Leopold 1968; Walsh et al. 2005). Urbanization and its subsequent increase in impervious surfaces leads to pollutants such as nutrients, sediments, metals, organic matter, hydrocarbons and oils impacting local waterways (Booth and Jackson 1997; Schueler 1994). Excess of these nutrients in surface water can cause eutrophication and algal blooms which has detrimental impacts on ecology and aquatic life (USEPA 1993). Regulations and stormwater technologies have continued to evolve in order to counteract the increase in water quality and quantity.

2.2 **Evolution of Stormwater Management Regulations**

Regulation of stormwater management dates back as early as 1760 BCE when King Hammurabi of Ancient Mesopotamia presented rules in the Code of Hammurabi to prevent flooding and protect downstream land owners from upstream farm practices. Section 53 of Hammurabi’s Code states that if a landowner does not keep up with repair or maintenance of ponds, berms or irrigation channels, and it floods downstream land, they are solely responsible for reimbursing for loss of crop or property (Echols and Pennypacker 2015).

Polluted runoff resulting from site development is subject to regulations that have been created and enacted over the last several decades at the federal, state and local levels. The Clean Water Act (CWA) was enacted in 1972 in order to ensure that the country’s waters remained fishable, swimmable and drinkable [33 USC 26 – Federal water pollution control act (1972)]. The National Pollutant Discharge Elimination System (NPDES) was originally included in the CWA to regulate point source discharges from industrial process wastewater and municipal sewage treatment. In 1987, Section 402(p) was added to the CWA, which redefined “point source” to include construction activity within Municipal Separate Storm Sewer Systems (MS4s) in order to create a permit system for stormwater runoff (USEPA 2005). Section 303d of the CWA also
establishes Total Maximum Daily Loads (TMDLs) for Waters of the United States, which guide stormwater management and restoration to maintain water quality parameters within a specified threshold (Clarke 2003). The Environmental Protection Agency (EPA) enforces these regulations on a federal level, but state and local governments can choose to enforce stricter thresholds. In order to meet water quality thresholds, any new development or redevelopment on a site has strict guidelines for implementing stormwater management practices.

2.3 EVOLUTION OF CURRENT STORMWATER MANAGEMENT

Traditionally, the goal of stormwater management was to provide flood control, with the priority being disposal of the water on-site as quickly as possible into receiving waters. As the negative impacts of discharging untreated stormwater into surface waters became apparent, the efforts transitioned to also include accounting for water quality in addition to quantity (Burian et al., 1999). In response to regulation, localities develop tailored Stormwater Management Programs (SMPs) that promote mimicking pre-development water quantity and quality through low impact development and best management practices (BMPs), also known as green-stormwater infrastructure (GSI), which provide runoff volume and quality control through physical, physiochemical and biological pollutant removal through smaller spatially-distributed facilities (Livingston et al. 1997). GSIs fall into a series of classifications by the method in which they contain and treat water, including infiltration, filtration, impoundment for gravitational settling, or a hybrid of these categories (USEPA 2005). For example, bioretention facilities are a common GSI that is installed on development projects. These are landscaped areas with a small ponding depth and an engineered porous soil media that physically filters sediment out of percolating water and also uses soil adsorption to extract other pollutants such as phosphorous and nitrogen. These facilities are specifically designed to the size and depth requirements to filter out a precise amount of pollutants based on the land use changes of the contributing drainage area (CDA), and are expected to maintain the designed pollutant removal rates throughout their lifetime (Davis et al. 2009). Other common GSI facilities include, permeable pavement, infiltration swales, sand filters and retention ponds.
2.4 THE IMPORTANCE OF MONITORING

An often-discussed controversy in stormwater management is the designed efficiency of GSI. Since these facilities are designed to remove a certain volume of pollutants, it is imperative that they meet their assigned removal efficiencies in practice, not only for downstream water quality but for budgetary concerns of the MS4 (Scholes et al. 2008). Although monitoring GSIs to verify the efficiency of the facility is important, it is not standard practice throughout industry. MS4s have the power to enforce monitoring programs at the local level; however, all current programs are voluntary and not legally mandated by federal or state regulation (Adair et al. 2014; Jones et al. 2004). Limited understanding of the physical and chemical processes occurring within these facilities has largely resulted in assignment of assumed long-term removal efficiencies which can have profound future environmental and economic consequences. A more complete understanding of the actual removal processes and long-term efficiencies through long-term standardized monitoring is required to refine design standards for future GSI.

2.5 EXISTING MONITORING PROTOCOLS

In order to effectively monitor GSI facilities, it is important to have a standard set of practices and methods that is constant for all facilities to produce and maintain a statistically significant set of data (Urbonas 1995). Existing evaluation programs include both interstate protocols such as Technology Acceptance Reciprocity Partnership (TARP 2003) and Technology Assessment Protocol – Ecology (TAPE 2011) and more localized protocols specifically for Virginia (VTAP 2012), Metropolitan North Georgia (NGTAP 2010), Nashville (NSWMM 2012), etc. The goal of each of these programs is to establish a uniform method of monitoring stormwater quality and to develop quality assurance across all aspects of the process. New technologies often face regulatory and financial road blocks that slow down implementation over a large scale or region. With a standard way to monitor and analyze the efficiency of a facility, it can be more quickly adopted and utilized in similar regions or municipalities.

2.5.1 Environmental Technology Verification Program (ETV)

The Environmental Technology Verification Program (ETV) was developed by the Environmental Protection Agency (EPA) is the only nationally established platform set up to evaluate and verify stormwater treatment technologies. It is not only limited to stormwater, but
also includes air, wastewater, and drinking water treatment, and is therefore broad in nature (Adair et al. 2014). It was created in 1995 and provided guidelines on monitoring studies to verify performance claims of innovative stormwater treatment technologies, therefore expanding the technology choices of public and private authorities (US EPA 2000). However, rather than a standard protocol, each study is customized for the needs of each setup and is therefore wide in scope, thereby providing diminished use for large-scale compilation and analysis of results from multiple installations (Sample et al. 2012). In 2007, the ETV program was a founding member of the International Working Group on Environmental Technology Verification (IWG-ETV), with other members including Canada and the European Union with the Philippines, South Korea, Japan and Denmark also joining and developing programs over time. Participating countries are working towards mutually recognize each other’s technology verifications in order to streamline technology approval and exemplify the program slogan, “Verify Once, Accept Everywhere” (EPA ETV 2016). The US ETV program transitioned from being funded by the EPA into a private vendor paid system; which put responsibility into the hands of treatment technology manufacturers, which without the backing of regulatory enforcement, led to a lack of interest in following through a verification process which caused the program to conclude in March 2014 (EPA ETV 2016).

2.5.2 Technology Acceptance Reciprocity Partnership (TARP)

TARP was created as a data sharing platform between six states (California, Massachusetts, Maryland, New Jersey, Pennsylvania, Virginia) and provided a uniform methodology to collect and analyze monitoring data (TARP 2003). This protocol aimed to provide a platform for interstate data sharing and reciprocity so that innovative treatment technologies were encouraged in practice and multiple entities could benefit from monitoring studies completed and performance verifications addressed. Despite the effort to promote consistent monitoring standards, there were several flaws with TARP that eventually led to its conclusion. One shortcoming of TARP is that the only required monitoring parameter is for sediment, whereas many other target pollutants emerging in impaired waters include nutrients, metals and oils (Sample et al. 2012). This protocol also was developed by states that are in opposite regions of the country with varying precipitation patterns and land characteristics. Prior to TARPs conclusion, California backed out of the program to pursue its own monitoring protocol that also
was suspended in development due to state budget constraints (Adair et al. 2014). With the disbanding of TARP, New Jersey and Massachusetts also released adaptations of the protocol as their standard monitoring protocols (NJCAT, MASTEP). Although the program now has officially disbanded, Illinois, Maryland and Pennsylvania, New Hampshire, Rhode Island, Ohio, New York as well as Denver, CO and Santa Monica, CA, accept efficiency verification of MTDs through TARP standards (Adair et al. 2014).

2.5.3 Technology Assessment Protocol – Ecology (TAPE)

One of the more widely accepted and utilized protocols is TAPE which was created in 2002 by the Washington State Department of Ecology (WSDOE) and backed by the Washington Stormwater Center (WSC). Due to budget and staffing constraints, the TAPE program was closed in 2008, but through an ecology grant and with the partnership of the City of Puyallup, Washington State University, University of Washington, and the Washington Stormwater Center, the program was revised and restarted in 2011 (WSC). The purpose of TAPE is to provide a regulatory certification process for emerging manufactured treatment devices (MTDs). The protocol helps identify a device’s effectiveness in removing pollutants in runoff and compares monitoring results with the performance claims of the proprietary device (TAPE 2011). Once approved and certified through the TAPE program, the treatment technology can be freely utilized throughout the state, which encourages implementation and development of innovative treatment technologies. Although the protocol was designed and is actively implemented in Washington State, many other states and localities have adopted the regulatory approval process for themselves. The TAPE Program addresses several pollutants including sediment, phosphorus, dissolved metals and oils but is limited in that it is designed for evaluating low volume, flow through MTDs (Sample et al. 2012; TAPE 2011). TAPE provides a Northwest region specific monitoring protocol for analyzing BMP efficiency, based on its high frequency low intensity rainfall trends, as indicated by storm event parameters including a 36 hour maximum duration.
2.5.4 Virginia Technology Assessment Protocol (VTAP) and Chesapeake Bay technology Assessment Protocol (CBTAP)

The Virginia Technology Assessment (VTAP) was developed in order to provide a methodology for monitoring MTDs to specifically address pollutant constituents that are applicable to stormwater design in Virginia that were commonly not addressed in existing protocols, including target pollutants such as phosphorus and nitrogen (Sample et al. 2012). Some noteworthy elements of VTAP include; a pilot, conditional and general use designation level system corresponds to an allowable number of sold units in an effort to fund field monitoring studies, requirement for a preliminary project plan to be submitted prior to field sampling in order to confirm study implementation, and also requiring sample analyses to be completed by certified laboratories. Due to funding issues by the Virginia Department of Environmental Quality (DEQ), the draft protocol was withdrawn. Efforts have been made by the Chesapeake Bay Program in order to adapt the draft protocol used for VTAP into a starting point for what is to be known as the Chesapeake Bay Technology Assessment Protocol (CBTAP). Similarly to VTAP, it is designed for monitoring MTDs and making them readily verified and freely utilized throughout the entire watershed once approved (“Concept for the Development of a Chesapeake Bay Technology Assessment Protocol (CBTAP)” 2014, D. Sample personal communication 2018).

2.6 NOTABLE MONITORING PARAMETERS

2.6.1 Minimum Number of Storms Sampled

In any monitoring program it is important to have assurance that samples collected are representative of facility conditions and that the quantity of storm events to be sampled are sufficient in order to provide statistically significant results. Although it is intuitive that more storm sampled means higher confidence in results, number of storms sampled are most often-times driven by budgetary constraints due to the inherently expensive nature of monitoring. Professional experience and best judgement loosely supplemented by methodology and equations often dictate minimum number of qualifying storm events to be sampled (Law et. al 2008). Burton and Pitt (Burton and Pitt 2002) presented an equation to estimate the number of samples required as a function of significance level, $\alpha$, statistical power, $\beta$, standard deviation as a fraction of the mean, COV, and allowable error, $\varepsilon$ (Equation 1).
Using typical values for the above parameters provides the number of samples in the 25-50 range ($\alpha = 0.05, \beta = 0.20, \varepsilon = 0.20$, COV ranging from 0.5 to 1.5 for common water quality parameters) (Pitt 2009). The allowable error term accounts for sampling, equipment, and laboratory error, which can be controlled to a degree with the implementation of detailed quality assurance and quality control procedures as well as equipment maintenance. This equation is referenced in several protocols including TAPE, VTAP, and CBTAP in order to determine sample size of storms to be monitored.

2.6.2 Inter-event Time

Minimum period without rainfall, also referred to as inter-event time is used to isolate an individual storm event from a long-term precipitation record. When statistically analyzing efficiency data for GSI facilities, all storm events are assumed to be independent in nature, meaning a previous storm event has no effect on the facilities removal efficiency during the following event. Storm characteristics such as total storm depth and rainfall intensity are influenced by the selection of a representative inter-event time. There is a wide range in the values adopted for inter-event time from just a few minutes to 24 hours which is all based on regions precipitation trends (high intensity, low duration, etc.) but the values of 6 to 8 hours have been widely adopted in many studies (Agnese et al. 2006; Asdak et al. 1998; Gyasi-Agyei & Melching 2012; Loukas & Quick 1996; Marin et al. 2000). In a study done by Driscoll (1989), precipitation data from three locations throughout the United States was analyzed for effect of inter-event dry time between precipitation events on other storm parameters such as depth, intensity and volume. This study found that although frequency of precipitation varies across the study areas (New Jersey, Texas and California), a 6-hour inter-event time is an appropriate minimum time for dry conditions between storm events, and should be used uniformly throughout the United States, to provide a consistent standard. Although no sources are stated in the reasoning for selection, an inter-event time of 6-hours is chosen by a large majority of the reviewed monitoring programs.
2.7 CONCLUSION

As development and urbanization increase, the need to effectively keep its impacts from degrading local waterways will also need to evolve. Stormwater regulations and technologies have evolved to address these issues, however there is a noticeable gap in literature addressing implementation of monitoring programs to test the intended and designed efficiency of GSI facilities. The use of GSI facilities is still relatively new and the long-term effects of implementation are not well-known, which is why the development of consistent methodology for verification is imperative to the continued evolution of stormwater management.
3. **Evaluation of Green Stormwater Infrastructure Monitoring Protocols**

3.1 **Introduction**

3.1.1 **Background**

Urbanization leads to impacts on watershed hydrology such as increased peak flows, increased runoff volumes, increased sediment loading and overall degradation of water quality (Leopold 1968; Walsh et al. 2005). Construction, and its subsequent increase in impervious surfaces, leads to pollutants such as nutrients, sediments, metals, organic matter, hydrocarbons and oils impacting local waterways (Booth and Jackson 1997; Schueler 1994). Due to these impacts, stormwater regulations have been enacted to protect the integrity of the nation’s waters.

In the United States, the first major legislative effort to address growing water quality impairments was the Refuse Act of 1899, a sub section of the River and Harbors Appropriation Act. This ultimately evolved into the Federal Water Pollution Control Act in 1948 and then into the more commonly referenced, Clean Water Act (CWA) of 1972. The CWA’s mandate is to ensure that the country’s waters remained fishable, swimmable and drinkable (33 USC 407, 33 USC 1251 respectively). The National Pollutant Discharge Elimination System (NPDES) was originally implemented as part of the CWA to regulate point source discharges from industrial process wastewater and municipal sewage treatment. In 1987, Section 402(p) was added to the CWA, which redefined “point source” to include construction activities within Municipal Separate Storm Sewer Systems (MS4s) which allowed permitting stormwater runoff from development projects (National Research Council 2009; USEPA 2005). In order to treat runoff to meet discharge standards established in response to regulatory thresholds, stormwater best management practices (BMPs) were developed to treat stormwater effluent prior to discharge into receiving channels/waters.

Current structural stormwater BMPs tend to mitigate quantity and quality of urban runoff on a spatially distributed, site scale. These BMPs are constructed to capture, treat and slowly release or infiltrate on-site flow, in the attempt to mimic pre-development conditions and thereby mitigate downstream problems such as sediment and nutrient loading in order to meet region specific total daily maximum loads (TMDLs) (Hopkins et al. 2017; Ice 2004; Livingston at al.
For the purpose of this paper, a distinction is made between BMPs and Green Stormwater Infrastructure (GSI). The term BMP encompasses all treatment facilities that mitigate stormwater quality, quantity, or a combination of the two. GSI focuses on the quality aspect of stormwater management and is collectively defined as storm water control measures (SCMs) that provide physical removal (e.g. infiltration, settling), as well as physio-chemical (e.g. adsorption), and biological (e.g. plant uptake) removal processes. GSI facilities include bioretention cells, permeable pavement, sand filters, vegetated swales, etc. Limited understanding of the physical and chemical processes occurring within GSI facilities has largely resulted in assignment of assumed long-term percent removal metrics, which have profound future environmental and economic consequences. A recent study has worked to quantify this uncertainty based on individual monitoring studies, revealing large error bars on commonly utilized GSI facilities (Aguilar & Dymond, under review). Sources of error in removal efficiencies can be spatial variability in soils and precipitation as well as the construction, operation, and maintenance of the facilities (National Research Council 2009; Simpson, Musgrove, & Korcak, 2004). The increased reliance on GSI facilities to comply with regional and local stormwater regulations begs the need for confirmation of facility efficiency through long-term monitoring programs.

In order to determine if installed GSI facilities are operating as claimed, long term consistent monitoring studies must be implemented. Although historic monitoring studies do exist, there are many variables that affect the monitored efficiency of a GSI facility. Irregular sampling and analysis methods used to estimate or eliminate these variables results in a wide degree of variability in reported field performances of various GSI facilities (Field et al. 2006; Geosyntec Consultants, 2009). Even if all variables are correctly accounted for, there is uncertainty in monitoring itself due to sampling error, equipment error, and lab error. One of the largest issues is the budgetary implications of conducting monitoring studies. Due to large scale issues such as the lack of regulatory backing for long-term scientific assessment, the expensive nature of conducting thorough monitoring studies precludes their wide-spread adoption. In an effort to curb overall cost and spread personnel and budgetary resources between multiple sites and facilities, a study in Philadelphia utilized a tiered approach that may be adaptable in other communities in which variable cost low, medium, and high-level monitoring is performed for
various facilities. High-level monitoring encompasses influent and effluent samplers, while low-level monitoring includes cheaper visual inspections occurring seasonally, with the assumption that certain surrogate parameters are indicative of data obtained in high-level monitoring (Welker et al. 2013). Monitoring through a standardized approach will aid in mitigating consistency issues and will assist in better understanding the processes (including removal efficiencies) at work in these facilities. Insight gained through monitoring will allow adoption of better regulations that are driven by consistent data collection and analysis.

3.1.2 Existing Protocols
Without the presence of a nationally recognized GSI monitoring protocol, states and cities have taken on the responsibility to address the need for efficiency verification in stormwater practices. Currently recognized active or inactive programs for GSI monitoring include widespread protocols such as Technology Acceptance Protocol – Ecology (TAPE 2011), Technology Acceptance Reciprocity Partnership (TARP 2003), Environmental Technology Verification Program (ETV) (EPA 2016), New Jersey Corporation for Advanced Technology (NJCAT), as well as smaller local protocols including Metropolitan North Georgia Technology Assessment Protocol (Post-Construction Stormwater Technology Assessment Protocol for Metropolitan North Georgia 2010) and Nashville Field Testing requirements (Metropolitan Nashville - Davidson County Stormwater Management Manual 2012). Some cities or states without their own tailored GSI monitoring protocol, have adapted one or more of these protocols as a means for approval of treatment technologies. For example, TAPE was created for, and is implemented by, Washington State but has reciprocity in New York, Maine, New Hampshire, Rhode Island and cities such as Denver, CO, St. Louis, MO, Portland, OR, and Sacramento, CA (TAPE 2011). A summary of program documents that were reviewed for this study are summarized in Table 1 and illustrated geographically in Figure 1. Note the current status of the programs and how overall, confirming BMP removal efficiencies through standardized monitoring programs have been attempted and proven successful in some areas but have a high chance of being discontinued due to lack of funding, applicability at a broader scale, and lack of regulatory backing.
<table>
<thead>
<tr>
<th>Program Name</th>
<th>Year</th>
<th>Place of Origin</th>
<th>Places that Recognize Program</th>
<th>Program Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology Acceptance Reciprocity (TARP)</td>
<td>2001</td>
<td>CA, NJ, MA, VA, PA, MD</td>
<td>States: NY, IL, OH, RI, NH, TX Cities: Portland, OR, Denver, CO, Santa Monica, CA</td>
<td>Inactive, but variations adapted by individual States</td>
</tr>
<tr>
<td>EPA Environmental Technology Verification (ETV) Program</td>
<td>1995</td>
<td>National</td>
<td>USA</td>
<td>Inactive</td>
</tr>
<tr>
<td>Virginia Technology Acceptance Protocol (VTAP)</td>
<td>2016</td>
<td>VA</td>
<td>-</td>
<td>Draft, discontinued</td>
</tr>
<tr>
<td>New Jersey Corporation for Advanced Technology (NJCAT)</td>
<td>2013</td>
<td>NJ</td>
<td>New Jersey</td>
<td>Active</td>
</tr>
<tr>
<td>Chesapeake Bay Technology Assessment Protocol (CBTAP)</td>
<td>2018</td>
<td>Chesapeake Bay</td>
<td>Chesapeake Bay</td>
<td>Draft</td>
</tr>
<tr>
<td>North Georgia Technology Acceptance Protocol (NGTAP)</td>
<td>2010</td>
<td>Atlanta, GA</td>
<td>North Georgia</td>
<td>Active</td>
</tr>
<tr>
<td>Massachusetts Stormwater Technology Evaluation Project (MASTEP)</td>
<td>2009</td>
<td>MA</td>
<td>Massachusetts</td>
<td>Inactive</td>
</tr>
<tr>
<td>Charlotte BMP Pilot Monitoring Program</td>
<td>1999</td>
<td>Charlotte, NC</td>
<td>Piedmont Region</td>
<td>Active</td>
</tr>
<tr>
<td>North Carolina Preliminary Evaluation Program (NCPEP)</td>
<td>1997</td>
<td>North Carolina</td>
<td>North Carolina</td>
<td>Active</td>
</tr>
<tr>
<td>Metropolitan Nashville - Field Testing Protocol</td>
<td>2012</td>
<td>Nashville, TN</td>
<td>Davidson County, TN</td>
<td>Active</td>
</tr>
</tbody>
</table>
Figure 3.1. Geographic location of all monitoring programs collected for this study, including place of origin as well as recognizing states or cities

Although most of the protocols reviewed were designed specifically for proprietary manufactured treatment devices, the protocols were included due to their transferability to non-proprietary GSI facilities such as bioretention cells, sand filters, permeable pavement, infiltration trenches, etc. Each of these protocols are similar in structure and purpose but vary in sampling procedures, frequency, required study parameters, and other details. A more detailed discussion of the primarily recognized protocols is included below. Despite resources to develop these programs, many of the protocols are either inactive or were discontinued in draft phase due to a lack of financial backing or other regulatory challenges, which raises concern about the ability and commitment to gather long term, consistent GSI monitoring data for the purpose of evaluating the true performance of these systems.
3.1.2.1 ETV

The Environmental Technology Verification Program (ETV) was developed by the Environmental Protection Agency (EPA), and is the only nationally established platform set up to evaluate and verify stormwater treatment technologies. Besides stormwater, the Program also includes air, wastewater, and drinking water treatment (Adair et al. 2014). It was created in 1995 and provided guidelines on monitoring studies to verify performance claims of innovative stormwater treatment technologies, therefore expanding the technology choices of public and private authorities (US EPA 2000). However, rather than a standard protocol, each study is customized for the needs of each setup, thus providing diminished use for large-scale compilation and analysis of results from multiple installations (Sample et al. 2012). In 2007, the ETV program was a founding member of the International Working Group on Environmental Technology Verification (IWG-ETV), with other members including Canada and the European Union with the Philippines, South Korea, Japan and Denmark also joining and developing programs over time. Participating countries are working to mutually recognize each other’s technology verifications in order to streamline technology approval and exemplify the program slogan, “Verify Once, Accept Everywhere” (EPA ETV 2016). The US ETV program transitioned from being funded by the EPA into a private vendor-paid system; which put responsibility into the hands of treatment technology manufacturers, which, without the backing of regulatory enforcement, led to a lack of interest in following through on a verification process. This caused the program to end in March 2014 (EPA ETV 2016).

3.1.2.2 TARP

TARP was created as a data sharing platform between six states (California, Massachusetts, Maryland, New Jersey, Pennsylvania, Virginia) and provided a uniform methodology to collect and analyze monitoring data (TARP 2003). This protocol aimed to provide a platform for interstate data sharing and reciprocity so that innovative treatment technologies were encouraged in practice and multiple entities could benefit from completed monitoring studies and performance verifications. Despite the effort to promote consistent monitoring standards, there were several flaws with TARP that eventually led to its conclusion. One shortcoming of TARP was that the only required monitoring parameter was for sediment, whereas many other target pollutants in impaired waters include nutrients, metals and oils (Sample et al. 2012). This
protocol also was developed by states that are in opposite regions of the country with varying precipitation patterns and land characteristics. Prior to TARP's conclusion, California backed out of the program to pursue its own monitoring protocol that also was suspended in development due to state budget constraints (Adair et al. 2014). With the disbanding of TARP, New Jersey and Massachusetts also released adaptations of the protocol as their standard monitoring protocols (MASTEP 2013; NJCAT 2006). Although the program now has officially disbanded, Illinois, Maryland and Pennsylvania, New Hampshire, Rhode Island, Ohio, New York as well as Denver, CO and Santa Monica, CA, accept efficiency verification of manufactured treatment devices (MTDs) through TARP standards (Adair et al. 2014).

3.1.2.3 TAPE
One of the more widely accepted and utilized protocols is TAPE which was created in 2002 by the Washington State Department of Ecology (WSDOE) and backed by the Washington Stormwater Center (WSC). Due to budget and staffing constraints, the TAPE program was closed in 2008, but through an ecology grant and with the partnership of the City of Puyallup, Washington State University, University of Washington, and the Washington Stormwater Center, the program was revised and restarted in 2011 (WSC). The purpose of TAPE is to provide a regulatory certification process for emerging MTDs. The protocol helps identify the effectiveness of devices in removing pollutants in runoff and compares monitoring results with the performance claims of the proprietary device (TAPE 2011). Once approved and certified through the TAPE program, the treatment technology can be freely utilized throughout the state, which encourages implementation and development of innovative treatment technologies. Although the protocol was designed and is actively implemented in Washington State, many other states and localities have adopted the regulatory approval process for themselves. The TAPE Program addresses several pollutants including sediment, phosphorus, dissolved metals and oils but is limited in that it is designed for evaluating low volume, flow through MTDs (Sample et al. 2012; TAPE 2011). TAPE provides a northwest region-specific monitoring protocol for analyzing BMP efficiency, based on its high frequency, low intensity rainfall trends, as indicated by storm event parameters, including a 36-hour maximum duration.
3.1.2.4 Smaller-Scale Protocols

Due to the lack of regulatory backing for development of standards and monitoring GSI facilities, states, cities and municipalities have attempted to develop and implement local monitoring guidelines. Virginia’s target pollutant is phosphorus, addressed more thoroughly in its proposed adapted protocol (Virginia Technology Assessment Protocol - VTAP); however, VTAP did not make it past draft form due to funding constraints (Sample et al. 2012; VTAP 2012). The Chesapeake Bay Program has revitalized efforts on VTAP but increased its scope to include the entire Chesapeake Bay in order to address overall nutrient issues. This new protocol will be released as the Chesapeake Bay Technology Assessment Protocol (CBTAP, unpublished) within the 2019 calendar year (Sample, personal communication, 2017). Other city-scale protocols have been implemented in an effort to address local needs, including the North Georgia Technology Assessment Protocol (NGTAP 2010), the Nashville, TN Stormwater Management Manual (NSWMM 2012), and the Charlotte Pilot BMP Monitoring Program (CPMP 2017). Local jurisdictions chose to create their own individual protocol due to lack of state or regional monitoring guidelines; however, this encourages a lack of collaboration with other cities.

3.1.2.5 Common Program Components

Regardless of scale or location of monitoring protocols review, all have similar components which include definition of monitoring parameters, storm event parameters, selection of monitoring equipment, identification of field and lab procedures during monitoring, and recommendations regarding lab and data analysis. Monitoring parameters includes items such as site selection and identification of constituents of interest. These typically vary based on region and can even vary intra-region depending on the types of impairments of local streams, often specified in local TMDL reports. Storm event parameters spans a range of factors including identification of minimum/maximum precipitation depths, definition of minimum inter-event time, maximum thresholds of storm duration or intensity, the minimum number of required storm events, or a specified minimum study length. Typically, these parameters are inter-related, and can be based on analysis of historic precipitation data. Selection of monitoring equipment is often site and BMP specific, although most installations include common types of equipment such as precipitation gauges, flow sensors, auto-sampling equipment, and data loggers/transmitters at a minimum. Identification of field and lab procedures as well as lab/data
analysis varies greatly across these reports. Most do not go into specifics with regards to actual lab procedures used for the isolation and quantification of the various constituents. Reports are consistent in stressing the importance of equipment and facility maintenance during the study period as well as quality control procedures to ensure that samples are consistently and cleanly handled during collection and transport to labs for analysis. An illustration of typical program development and implementation is shown in Figure 2.

![Monitoring Program Development](image)

**Figure 3.2. Flowchart depicting typical items associated with creating a monitoring program**

### 3.1.3 Purpose

While the typical components necessary for implementation of monitoring plans is consistent across protocols, there has been little guidance regarding adaptation of historic protocols for specific regional application. Many of the components, such as equipment, lab analysis procedures, and data analysis are standard and are directly translatable across large regions. However, development of storm event parameters should include a detailed analysis of local data to design the most comprehensive monitoring plan. A methodology to identify essential components of a monitoring protocols and determine monitoring and sampling frequencies is
necessary to design consistent translatable monitoring efforts that will aid the scientific community in more fully understanding the functionality of GSI. The purpose of this study is to identify essential components of a monitoring protocol for non-proprietary stormwater facilities and provide a methodology for analysis of precipitation data for identification of many protocol parameters.

3.2 METHODS

3.2.1 Analysis of Existing Protocol Components
Several spatially diverse monitoring protocols were reviewed to synthesize current methodologies for developing a monitoring study. This study focused on extracting data (see Tables 2 and 3) from any manuals or documents that dictate sampling methodology and/or data analysis and efficiency calculations for stormwater facilities, without regard to the type of facility. While most protocols found were designed specifically for proprietary manufactured treatment devices, it is believed that many of the monitoring parameters are transferable to studies related to GSI facilities such as bioretention cells, sand filters, permeable pavement, infiltration trenches, etc.

Selected documents were reviewed for both quantitative (Table 2) and qualitative (Table 3) information. Quantitative variables include details pertinent to collecting samples that would be indicative of removal processes through a GSI facility including monitoring study duration, minimum storm depth required, minimum inter-event dry period between monitored storms, minimum number of sampled storm events, etc. Qualitative information provides insight about the necessary pieces of the monitoring program such as minimum equipment requirements, types of flow sampled, what pollutant constituents are to be sampled, presence of procedures for sample collection techniques, equipment maintenance, data management and quality assurance and quality checks in the field and laboratory. The qualitative information was collected in a dichotomous manner where a “yes” or “no” corresponds to 1 and 0, respectively.

It is important to note that no statistical comparison analysis was performed on this data due to several inherent issues which could not be overcome. These include: small sample size of protocols, some smaller city-scale protocols reference broader protocols in regard to sampling
frequency and methodology creating a doubling effect, and the inactivity and draft designation for several protocols. Despite the lack of a statistical comparison, a general review of parameters, sampling frequencies and necessary included procedures is discussed in the study results and discussion section of this paper.

Table 3.2. Quantitative Comparison of Existing Protocols. NCPEP and MASTEP were not included in this table due to their lack of protocol documentation available.

<table>
<thead>
<tr>
<th>Program Name</th>
<th>Storm Event</th>
<th>Sampling</th>
<th>Monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min. Depth (cm)</td>
<td>Min. Inter-event Dry Period (hrs)</td>
<td>Min. Duration (hrs)</td>
</tr>
<tr>
<td>ETV*</td>
<td>0.51</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TAPE*</td>
<td>0.38</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>TARP*</td>
<td>0.25</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>VTAP</td>
<td>0.25</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>CBTAP</td>
<td>0.25</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>NCPEP</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NJCAT*</td>
<td>0.25</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>NGTAP</td>
<td>0.38</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Nashville</td>
<td>0.25</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>Charlotte</td>
<td>0.25</td>
<td>72</td>
<td>-</td>
</tr>
</tbody>
</table>

*Widely-adopted protocols
Table 3.3. Qualitative Comparison Summary Table of Existing Protocols: Type of facility – “N”: non-proprietary, “P”: proprietary, “B”: both. An entry of 1 and 0 correspond to a yes or no, respectively. NCPEP and MASTEP were not included in this table due to their lack of protocol documentation available.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TAPE</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TARP</td>
<td>N</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ETV</td>
<td>B</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VTAP</td>
<td>P</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CBTAP</td>
<td>P</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NJCAT</td>
<td>N</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NGTAP</td>
<td>P</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nashville</td>
<td>P</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Charlotte</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
3.2.2 Historic Precipitation Analysis

Measurement of stormwater constituents is dependent on runoff to dislodge and transport contaminants to downstream locations; therefore, it is precipitation dependent. Because of this, many components of a monitoring plan are related to various metrics regarding local precipitation during the study period. Various sources of historic data may exist, including National Oceanic and Atmospheric Administration (NOAA), United States Geologic Service (USGS), or local gauge networks. Analysis of gauging data from these locations typically requires the following steps:

1) Data download and verification of complete record
2) Adjustment of raw data, as necessary for iterative analysis
3) Creation of automated computational procedures for sliding window (storm event) data analysis
4) Review of results through data analysis
5) Recommendation of precipitation-based monitoring parameters

General items listed above are dependent on the data-set provided for analysis. For instance, data downloaded from NOAA gauges typically is cumulative within a one-hour duration before resetting on subsequent hours. Full understanding for data format is crucial in this case to prevent overestimating of storm totals and intensities. For this study, Visual Basic for Applications (VBA) automation within Microsoft Excel was used for data analysis due to the ease of processing collected data and reporting tabulated results. However, many other packages such as R, Matlab, etc. could also be used to produce similar results.

Raw data was processed to remove all inter-event times that had readings of zero precipitation. This processing reduced the size of data files and simplified the identification of storm start and stop time for developed VBA macros. The developed macros allowed input of user specified inter-event time (time with no precipitation) and required minimum cumulative storm rainfall. From this input, the VBA macro analyzed annual data to determine parameters including number of qualifying events, times of storm start/stop, total annual rainfall captured by qualifying events, total storm event precipitation, average storm intensity, and peak storm intensity. This data can
then be used to select appropriate parameters (matching those found in Table 2) based on historic precipitation data.

3.3 **CASE STUDY: DEVELOPMENT OF MONITORING PROTOCOL FOR FAIRFAX COUNTY, VIRGINIA**

The methodology developed in the preceding section was applied in Fairfax County, Virginia to create a proposed monitoring protocol for use in the analysis of non-proprietary GSI facilities. Fairfax County, Virginia is located at the mouth of the Chesapeake Bay, where strict TMDLs are enforced for upstream municipalities. Fairfax County allocates a large portion of their water resources funds towards ensuring the operation and maintenance of thousands of GSI facilities. The integrity of local waterways directly depends on the longevity and efficiency of upstream GSI facilities. In addition to investing in GSI facilities operation and maintenance, a large portion of their budget also goes towards stream restoration of local waterways. The County has previously monitored individual facilities, but results have been inconsistent due to the lack of a unifying tailored protocol based on past precipitation data and a synthesis and refinement of current regional or national practices. The benefit of developing these procedures within the County includes standardized practices and methods for sampling, data analysis and documentation to determine facility performance county- and perhaps region-wide.

3.4 **REQUIRED MONITORING COMPONENTS**

Monitoring program components as identified in Table 3 were evaluated for inclusion in the protocol developed for Fairfax County. The nationally based ETV, TAPE, and TARP were given higher weight (2x) in considering applicable components due to their wide-scale applicability and history of use. In addition, these protocols are applicable to non-proprietary devices based on published guidance. The CBTAP (in development) document was also given higher weight in considering required program components since it is being developed for the Chesapeake Bay region, which encompasses all of Fairfax County. As mentioned previously, inclusion of the requirements in Table 3 are dichotomous in nature, but values from all protocols were summed (including double weighting of four protocols listed previously) and divided by 9 (# of protocols). In general, resulting values that were greater than or equal to 0.50 were considered desirable protocol features. Further consideration of low scoring items resulted in the
omission of several additional parameters for the recommended Fairfax County protocol. Recommended items are too numerous to discuss in detail in this document; however, the list of excluded items is 1) required discrete flow sampling, 2) a required health and safety plan beyond normal County procedures, 3) requirement for secondary flow device, 4) requirement for by-pass sampling. Although items 3 and 4 did score slightly above the 0.5 threshold (0.67 and 0.56, respectively), further consideration with respect Fairfax County’s monitoring goals resulted in omission of these relatively low scoring candidate components.

3.5 STORM EVENT COMPONENTS

Evaluation of storm event components was performed using spatially distributed historic rain gauge data provided by the County. This data consisted of 15-minute precipitation continuous precipitation data ranging from 1996 to 2016 for 10 rain gauges within the County (Figure 3). Data from 1996-2003 and 2011 was determined to be incomplete, with missing gaps in months or entire sets of gauge data missing and was therefore not used. Processing for supplied data required an initial assumption for inter-event time and minimum rainfall depth. Inter-event time is indicative of the effect of a preceding storm on a current storm. When statistically analyzing efficiency data for GSI facilities, all storm events are assumed to be independent in nature, meaning a previous storm event has no effect on the facilities removal efficiency, therefore, establishing an inter-event time that effectively separates the impact of a previous storm from a current storm is critical. The inter-event time between qualifying storm events was established to be 6-hours based on Driscoll et al., (1989) as well as the consensus of all reviewed monitoring protocols.
Minimum study duration was established as two years due to the annual variability in precipitation and climate data. Determination of the number of sampled storm events to adequately represent a facility’s performance is one of the most difficult parts of developing a monitoring protocol. A power analysis can be used to determine an ideal sample size based on input such as mean, standard deviation, but those are not often available before a study occurs. Burton and Pitt (Burton and Pitt 2002) presented an equation to estimate the number of samples required as a function of significance level, \( \alpha \), statistical power, \( \beta \), standard deviation as a fraction of the mean, COV, and allowable error, \( \varepsilon \) (Equation 1).

\[
    n = \left[ \frac{COV(Z_{1-\alpha}+Z_{1-\beta})}{\varepsilon} \right]^2
\]

(Equation 1)

Using typical values for the above parameters provides the number of samples in the 25-50 range \( (\alpha = 0.05, \beta = 0.20, \varepsilon = 0.20, \text{COV ranging from 0.5 to 1.5 for common water quality parameters}) \) (Pitt 2009). The allowable error term accounts for sampling, equipment, and laboratory error, which can be controlled to a degree with the implementation of detailed quality assurance and quality control procedures as well as equipment maintenance. Considering the budgetary implications of more sampling events, 25 storm events per study was chosen to be monitored over a minimum 2-year study duration.
Initially, a VBA analysis was run to identify the number of candidate storms based on a minimum storm depth of 0.25 cm (0.1 inches), as specified by the guidelines of most reviewed protocols. It is worth noting that this minimum depth requirement is specified largely based on protocols developed for proprietary storm water treatment devices characterized as low volume flow-through facilities with smaller contributing watersheds. Using these constraints, a frequency table of the number of qualifying storm events per year was created (Figure 4). On average, there were 57 qualifying storm events occurring annually. Although 57 storm events qualified, this is not a realistic number of storms that can be sampled. An approximate 20% decrease in monitored events due to sampling error, contamination or equipment failure (Law et al. 2008), should be applied to estimate the final number of monitored storms that will be available for analysis. Although a larger number of qualifying storm events creates a larger sample pool and therefore higher confidence in results, sampling, collecting and analyzing data from that many events would cause budgetary and personnel constraints. Storms qualifying with 6-hour inter-event time and a minimum storm depth of 0.25 cm (0.1 inches), on average contribute to approximately 96% of the annual precipitation in the County. While it is important to quantify and monitor a representative percentage of annual rainfall, which is referenced in some protocols as greater than 50% (Nashville 2012, NJCAT, TARP, 2003), it is also important to consider budgetary constraints.

![Figure 3.4. Average number of qualifying storm events for consolidated data from all gaging stations in Fairfax County, VA based on a minimum 0.25 cm storm depth and 6 hour inter-event time](image-url)
Measurable outflow is required to effectively monitor flow volumes and effluent pollutant concentrations from GSI facilities such as bioretention cells, infiltration trenches, and dry swales. To account for more than 50% of annual precipitation, while also reserving a 20% failure rate of collection, the ideal minimum storm depth requirement was found to be 1.27 cm for Fairfax County, VA. For the years of data analyzed, there was an average of 25 qualifying storm events per year, with individual years shown in Figure 5. The number of qualifying storm events per year is 25 and the minimum number of required sampled storm events is also 25. These 25 sampled storm events can occur over a recommended minimum two-year monitoring study period and should be used as a minimum value due to the increase in confidence in results as sample size increases. Ultimately, necessary values for the storm event components were determined using local County precipitation data with a 6-hour inter-event time and 1.27 cm (0.5 inch) minimum storm depth. Resulting values are shown in Table 4.
Figure 3.5. Average number of qualifying storm events for consolidated data from all gaging stations in Fairfax County, VA based on a minimum 1.27 cm storm depth and 6 hour inter-event time

Table 3.4. Summary Table of Recommended Storm Event Parameters for Fairfax County, VA

<table>
<thead>
<tr>
<th>Program Name</th>
<th>Storm Event</th>
<th>Sampling</th>
<th>Monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min. Depth (cm)</td>
<td>Min. Inter-event Dry Period (hrs)</td>
<td>Max Duration (hrs)</td>
</tr>
<tr>
<td>Fairfax</td>
<td>1.27</td>
<td>6</td>
<td>--</td>
</tr>
</tbody>
</table>
3.6 RESULTS AND DISCUSSION

The methodology discussed above was applied to a case study region to determine applicability of using past protocol recommendations supplemented with applicable discussion in the literature for development of a new protocol. Ultimately, the effectiveness of this procedure cannot be evaluated until applied to GSI monitoring studies by Fairfax County. Despite that constraint, several results of the applied methodology and resulting protocol can be discussed.

3.6.1 Assignment of Protocol Components by Weighted Dichotomous Data Scoring

Assignment of weights appeared to effectively select components from past protocols that were desired to be part of the Fairfax County protocol. However, most protocols required the majority of the components; therefore, making the scoring system irrelevant for these parameters’ inclusion in the final protocol. In addition, the low scoring items that were ultimately removed from the protocol were typically more expensive and therefore could have been deselected from past studies based on economic bias instead of scientific merit. Therefore, it is recommended that the list of program components in Table 3 be considered while taking into account all economic restrictions and scientific goals of the planned program. When considered from this perspective, the scoring system may be irrelevant; although, this did produce scores that aligned well with the desired program desires of the County.

3.6.2 Application of Automated Local Precipitation Analysis

Integration of local precipitation data in development of specific precipitation-based program components through automated VBA macros streamlined the decision-making process. As noted above, the initial minimum required cumulative storm depth based on already developed protocols is 0.25 cm. Ultimately, it was determined that this resulted in an excessive amount of storms within the County for the desired number of storm events for analysis. This is notable, since the selected 1.27 cm rainfall depth is much more likely to produce measurable effluent in the targeted GSI facilities than the 0.25 cm value. It is surmised that the 0.25 cm value recommended by most past protocols is due to their primary focus on proprietary BMPs which typically use some sort of mechanical or hydrodynamic stripping instead of media-based filtration. Typically, proprietary devices offer virtually no runoff reduction capability, while GSI
facilities give significant storage and uptake capabilities, therefore requiring higher cumulative rainfalls to produce measurable effluent.

The inter-event time is based both on the common value found in past protocols and the observations of results from other trial inter-event times using the VBA analysis. As described previously, for a 6 hour inter-event time (with 1.27 cm of minimum total storm rainfall), an average of 25 storms per year was generated. A 12-hour inter-event duration also produces 25 storms per year; however, the durations of these ‘storms’ is much longer and may actually consist of multiple rapid succession discrete events (such as summer thunderstorms). This trend is intensified when the inter-event time is increased further (to 36 or 48 hours), resulting in multi-day ‘storms’ which do not align with true climate conditions. These issues would seem to indicate that longer inter-event times are not recommended and that the lower 6-hour inter-event time as established in past protocols is preferred.

Note that no maximum storm intensity was specified for the Fairfax County case study. Typically, the maximum storm intensities occur during summer thunderstorms that produce significant runoff over short period. These intense events are typically very effective in producing the energy necessary to dislodge and convey sediment and bound pollutants to treatment locations. Therefore, they are an integral component of the monitoring effort. However, if the intensity is so high that the facility is overwhelmed and the overflow device (spillway or riser) is activated, then the sample may need to be eliminated from the study. This decision should be on a storm-by-storm basis and also based on the configuration of the outflow monitoring site location.

3.7 CONCLUSION

The importance of GSI monitoring has slowly been gaining recognition as a vital step in confirming facility performance and ensuring the health of local waterways. GSI performance and resulting monitoring practices vary based on precipitation patterns and watershed characteristics leading to difficulties adapting a uniform approach. Protocols for SCM evaluation have been established at the national, state and local levels, all to address the common issue of verifying performance efficiency. The reviewed protocol program documents range in detail,
scope and effectiveness, with some programs well-established, accepted and referenced while others struggle to sustain or have become inactive due to lack of enforcement or budgetary issues. Developing and implementing a monitoring program requires budgetary resources and time. An overarching national program that specifies field and lab details with the ability to adapt certain parameters specifically at the state level or small-region scale would be ideal in order to maximize implementation success is highly recommended. While this study did not reach that level, it does outline necessary components and procedures to use past protocols and local precipitation data in the development of a local or regional monitoring protocol.

Until the time that a national framework is in place for enforcing verifications on performance claims and providing standardized monitoring practices, state or small regions should consider the benefit of implementing a framework containing key elements discussed in this study. Collecting data on current GSI in a scientifically consistent manner will add valuable insight on these facilities. This insight can be used to gain a better understanding into the treatment mechanisms and efficiencies of these facilities and provide direction on the most efficient means to maintain GSI, thereby prolonging the lifespan of these facilities.
4. **CONCLUSION**

4.1 **IMPLICATIONS**

The increased implementation of GSI facilities to counteract impacts to surface waters due to urbanized development begs the need for confirmation of facility efficiency through consistent long-term monitoring programs. Due to the lack of a nationally recognized standard and regulation for confirming and verifying removal efficiencies of these facilities, monitoring should be implemented by state or small-scale regions using the specific considerations mentioned. Collecting data on installed GSI facilities in a scientifically consistent manner will add valuable insight into the efficiency and continued reliance on these facilities. This insight can be used to gain a better understanding into the treatment mechanisms and efficiencies of these facilities and provide direction on the most effective practices and timelines to maintaining GSI facilities, thereby prolonging lifespan.

A handful of protocols for SCM evaluation have been established at the national, state and local levels, all to address the common issue of verifying performance efficiency. The reviewed protocol program documents range in detail, scope, and effectiveness, with some programs well-established, accepted and referenced while others struggle to sustain or have become inactive due to lack of enforcement or budgetary issues. The methodology developed and utilized in the case study for Fairfax County can be a useful resource for other small-regions and municipalities in implementing GSI monitoring programs with tailored sampling and monitoring frequencies for parameters such as minimum storm depth, inter-event time and minimum number of sampled storm events.

4.2 **FUTURE WORK**

This study developed monitoring and sampling frequencies and parameters specific for Fairfax County, Virginia. Fairfax County precipitation trends indicated storm with high intensities and short duration with relatively short period of dry time between events. Future research should be conducted in utilizing the developed approach in other regions of varying precipitation trends in order to assess the transferability of methodology used in this study. Additionally, by providing Fairfax County with the tools to implement a monitoring program across various GSI facilities,
future research can be conducted in order to assess the overall effectiveness of the developed program based on monitoring studies completed within the County.

This study also revealed the inherent need for a national standardized and enforceable framework for confirming facility performance claims, which would be beneficial towards the evolution of stormwater management. Although field and lab procedures should remain the same, due to varying rainfall characteristics, an all-purpose set of sampling and monitoring frequencies cannot be used to cover the entire contiguous United States. Monitoring frequencies will vary based on the relationship of annual rainfall and storm intensity. Research can be developed to investigate the feasibility of a national verification program with a state or local jurisdiction certification; meaning that the verification protocol can cover all aspects of monitoring with local certification programs specifying any adjustments necessary. US Climate regions developed by the National Oceanic and Atmospheric Administration (NOAA) (Karl and Koss 1984) could be potentially useful to divide regions based on climate characteristics with adapted monitoring and sampling frequencies. EPA Ecoregions can also be considered due their distinction of regions based on geology, soils, vegetation, climate, land use and hydrology (Omernik 1987). The creation of a regulated, standard, and succinct methodology of monitoring GSI facilities will help advance technology and provide confidence in implementation.
5. REFERENCES


NJCAT Technology Verification Program. (n.d.). Retrieved from
http://www.njcat.org/verification-process/general-protocol.html


https://doi.org/10.1111/j.1467-8306.1987.tb00149.x


https://doi.org/10.1061/(ASCE)0733-9496(1995)121:1(23)


USEPA. (2005). National Management Measures to Control Nonpoint Source Pollution from

